2020 Building Performance Analysis Conference and SimBuild
co-organized by ASHRAE and IBPSA-USA

Conference Paper Session 16 Urban Scale Modeling and Working with Big Data

Joshua New, Ph.D.,
CEM, PMP, CMVP, CSM, IREE
Oak Ridge National Laboratory
newjr@ornl.gov

Urban-scale Energy Modeling: Scaling Beyond Tax Assessor Data

Learning Objectives

- Explain limitations to scalability of tax assessor data.
- Identify sources of data for urban-scale building energy modeling.
- Distinguish strengths/weaknesses of a dataset by defining comparative matrix fields.

ASHRAE is a Registered Provider with The American Institute of Architects Continuing Education Systems. Credit earned on completion of this program will be reported to ASHRAE Records for AIA members. Certificates of Completion for non-AIA members are available on request.

This program is registered with the AIA/ASHRAE for continuing professional education. As such, it does not include content that may be deemed or construed to be an approval or endorsement by the AIA of any material of construction or any method or manner of handling, using, distributing, or dealing in any material or product. Questions related to specific materials, methods, and services will be addressed at the conclusion of this presentation.

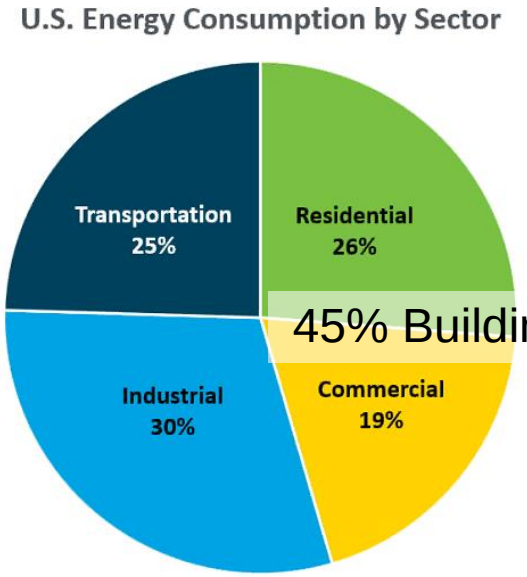
Acknowledgements

- Funding
  - U.S. Department of Energy
  - National Nuclear Security Administration
  - Building Technologies Office
  - Office of Electricity
  - Oak Ridge National Laboratory
  - Oak Ridge Leadership Computing Facility – HPC1
  - Argonne Leadership Computing Facility – HPC2
- Team Members
  - ORNL – Jiangye Yuan, Hsiuhan (Lexie) Yang, Jibonananda (Jibo) Sanyal, Olufemi (Femi) Omitaomu
  - EPB – William (Bill) Copeland, James (Jim) Ingraham, Lilian Bruce
  - Counties of Hamilton, Rhea, and Marion
- Software
  - EnergyPlus – Software 1
  - OpenStudio – Software 2
  - Automatic Building Energy Modeling (AutoBEM) – Software 3
    - AutoGen – Software 4
    - AutoSim – Software 5

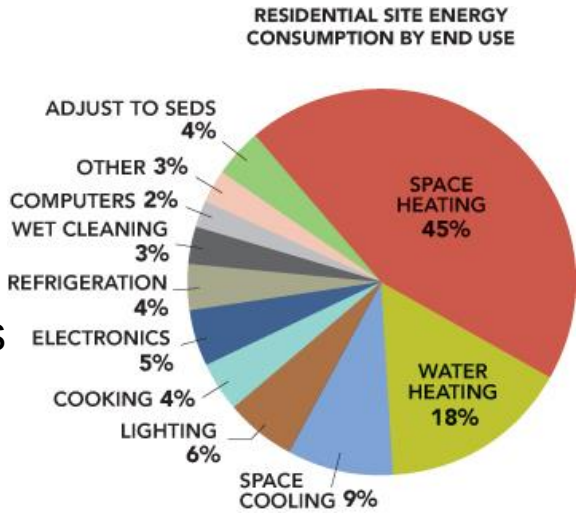
Outline/Agenda

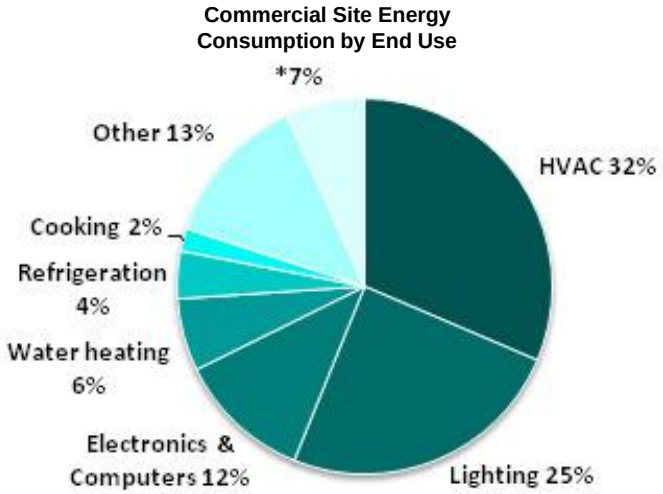
- Context
- Tax Assessor's Data
- Comparison Matrix
- Scalable Data Sources

Context



45% Buildings





Buildings consume 73% of the nation's electricity

Source: U.S. Energy Information Administration, January 2016 to January 2017, *Monthly Energy Review* – Table 2.1.

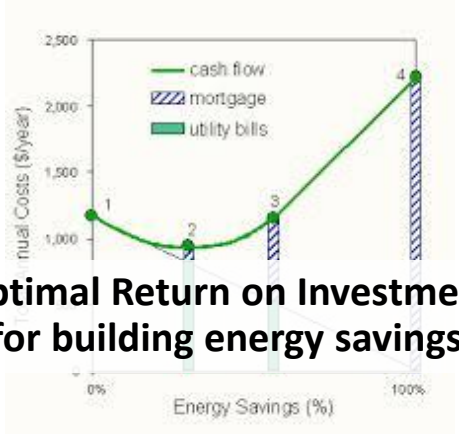
125 million U.S. buildings
$412 billion/yr energy bills (2019)

Goal of the DOE
Building Technologies Office:
30% energy reduction per sq. ft.
by 2030 compared to 2010 baseline

Building Energy Modeling – building descriptions + weather = estimated building energy consumption

$9B/yr – ESCO; $7B/yr – utility EE
$14B/yr – DR management systems
0.3% modified, BEM < 10% of those

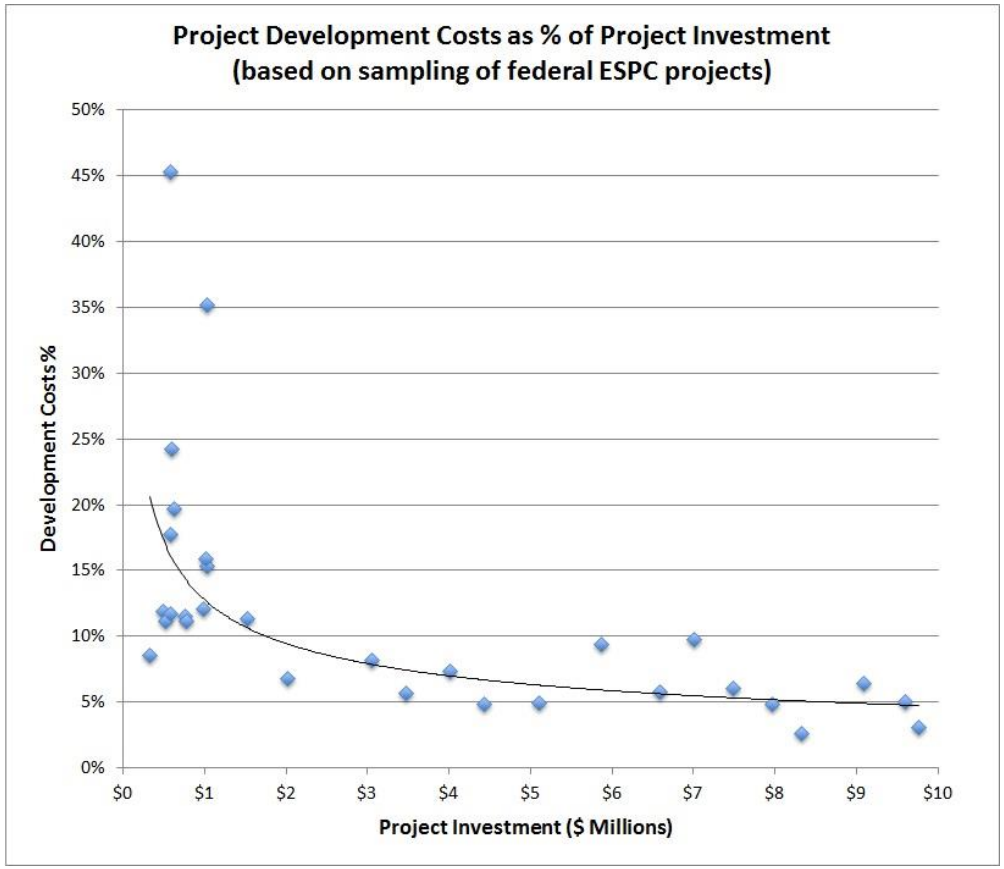
Context – Software 1 & 2



Optimal Return on Investment
(for building energy savings)

Simulation Engine and Analysis Platform
U.S. Dept. of Energy
$93M, 1995–?

Free, open-source (GitHub),
free support community (unmethours.com)



Context – HPC 1 & 2, Software 5

- World's fastest buildings energy model (BEM) simulator
- >500k building simulations in <1 hour
- 130M US buildings could be simulated in 2 weeks
- 8M simulations of DOE prototypes (270 TB)

19M core-hours
June 2, 2020

HPC1 Titan

CPU Cores	Wall-clock Time (mm:ss)	Data Size	EnergyPlus Simulations
16	18:14	5 GB	64
32	18:19	11 GB	128
64	18:34	22 GB	256
128	18:22	44 GB	512
256	20:30	88 GB	1,024
16,384	26:11	5.6 TB	65,536
32,768	31:29	11.5 TB	131,072
65,536	44:52	23 TB	262,144
131,072	68:08	45 TB	524,288

HPC2 Theta

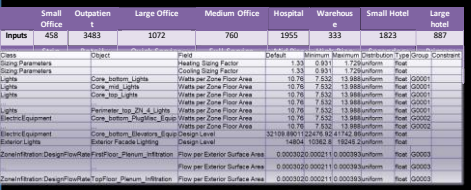
CPU Cores	Wall-clock Time (mm:ss)	Data Size	EnergyPlus Simulations
57,344	20:44	440 GB	229,376
114,688	28:20	880 GB	458,752

Context – vision

Digital Twin of every U.S. building

Methodology: Scalable compute, data, simulation, and empirical validation

1. Quantitatively rank most important building inputs



	Small Office	Outpatien t	Large Office	Medium Office	Hospital	Warehous e	Small Hotel	Large hotel
Inputs	458	3483	1072	760	1955	333	1823	887

Class	Object	Field	Default	Minimum	Maximum	Distribution	Type	Group	Constraint
Sizing Parameters		Heating Sizing Factor	1.33	0.931	1.729	uniform	float		
Sizing Parameters		Cooling Sizing Factor	1.33	0.931	1.729	uniform	float		
Lights	Core_bottom_Lights	Watts per Zone Floor Area	10.76	7.532	13.988	uniform	float	G0001	
Lights	Core_mid_Lights	Watts per Zone Floor Area	10.76	7.532	13.988	uniform	float	G0001	
Lights	Core_top_Lights	Watts per Zone Floor Area	10.76	7.532	13.988	uniform	float	G0001	
		Watts per Zone Floor Area	10.76	7.532	13.988	uniform	float	G0001	
Lights	Perimeter_top_ZN_4_Lights	Watts per Zone Floor Area	10.76	7.532	13.988	uniform	float	G0001	
ElectricEquipment	Core_bottom_PlugElec_Equip	Watts per Zone Floor Area	10.76	7.532	13.988	uniform	float	G0002	
		Watts per Zone Floor Area	10.76	7.532	13.988	uniform	float	G0002	
ElectricEquipment	Core_bottom_Elevators_Equip	Design Level	32109.89013	22476.92	41742.86	uniform	float		
Exterior Lights	Exterior Facade Lighting	Design Level	14804	10362.8	19246.2	uniform	float		
ZoneInfiltration DesignFlowRate	FirstFloor_Plenum_Infiltration	Flow per Exterior Surface Area	0.000302	0.000211	0.000393	uniform	float	G0003	
		Flow per Exterior Surface Area	0.000302	0.000211	0.000393	uniform	float	G0003	

Sensitivity Analysis

2. Time on world's #1 fastest high-performance machines

Supercomputers

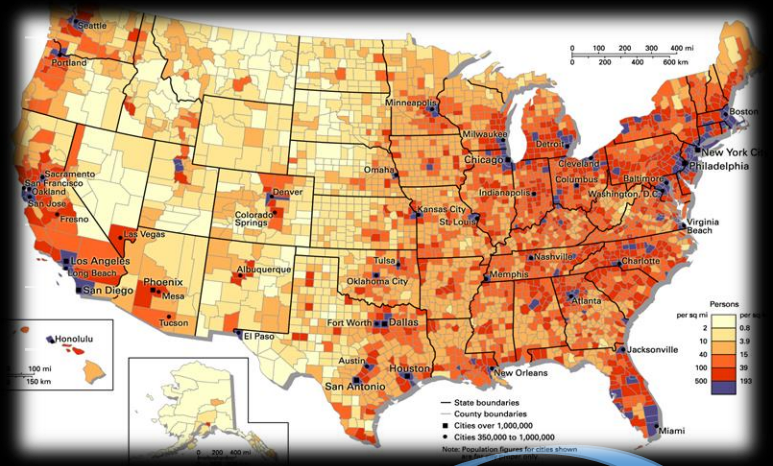
3. Identify and compare data sources for important inputs

	Short Title
Summary	Satellite imagery, including panchromatic and multispectral images
Data type	Image
Company	
Website	
Temporal resolution	Cities - 3-11 times per week
Spatial resolution	0.3 m
Measure accuracy	
Cost	$11 per sq. km
Format	GeoTiff
Mapping to building input variables	Building footprints
Mapping to area properties	Vegetated areas, road surface, buildings, parking lots
Mapping to material properties	Road pavement materials (e.g., concrete, asphalt), parking lots (e.g., gravel, soil)
Coverage of US	Over 10 million km² of coverage of the contiguous US
Orientation	Aerial
Existing internal software	N/A
Existing expertise	Remote sensing data analysis tool
Restrictions	N/A
Comments	

Comparison Matrix

4. Establish partnerships and APIs for scalable data retrieval

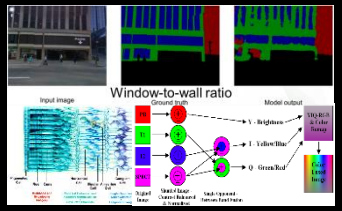
Databases



Demonstrate and stimulate GEB opportunities toward a sustainable built environment

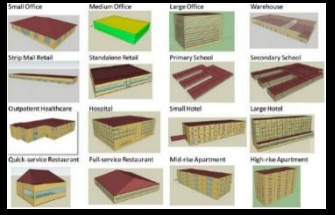
HPC2, Software 4-5, Software 1-2

5. Algorithms to extract building properties



Computer Vision

6. Create BEM data and models



DOE Prototype Buildings

7. Make BEM info freely available online

Download BEM via street address

Use cases:

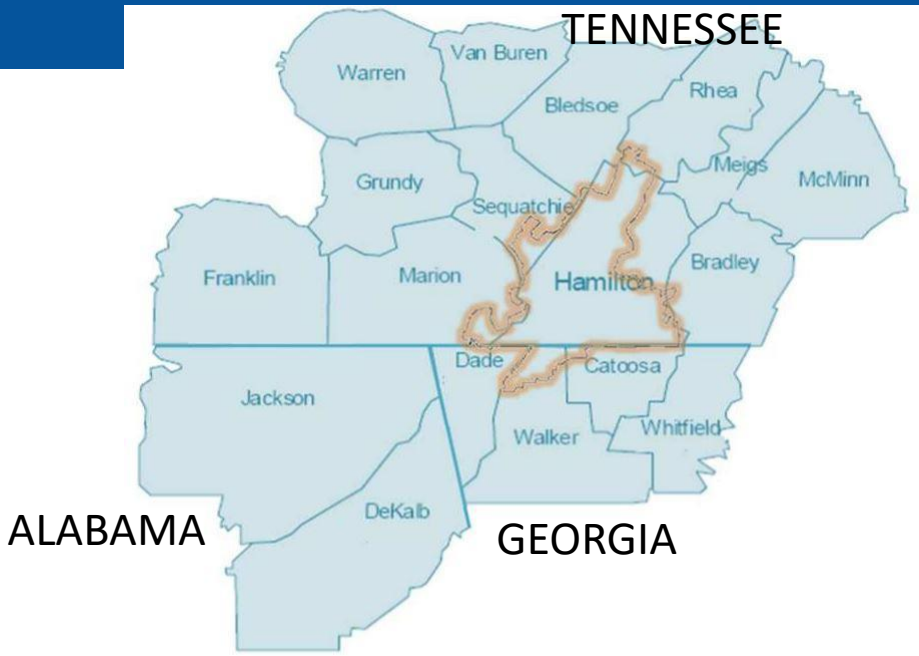
- Simulation-informed analysis
- Sales/market leads
- Utility program formulation
- Rate structures
- Resilience
- Automated financing
- Business model evaluation

Outline/Agenda

- Context
- Tax Assessor's Data
- Comparison Matrix
- Scalable Data Sources

Tax data

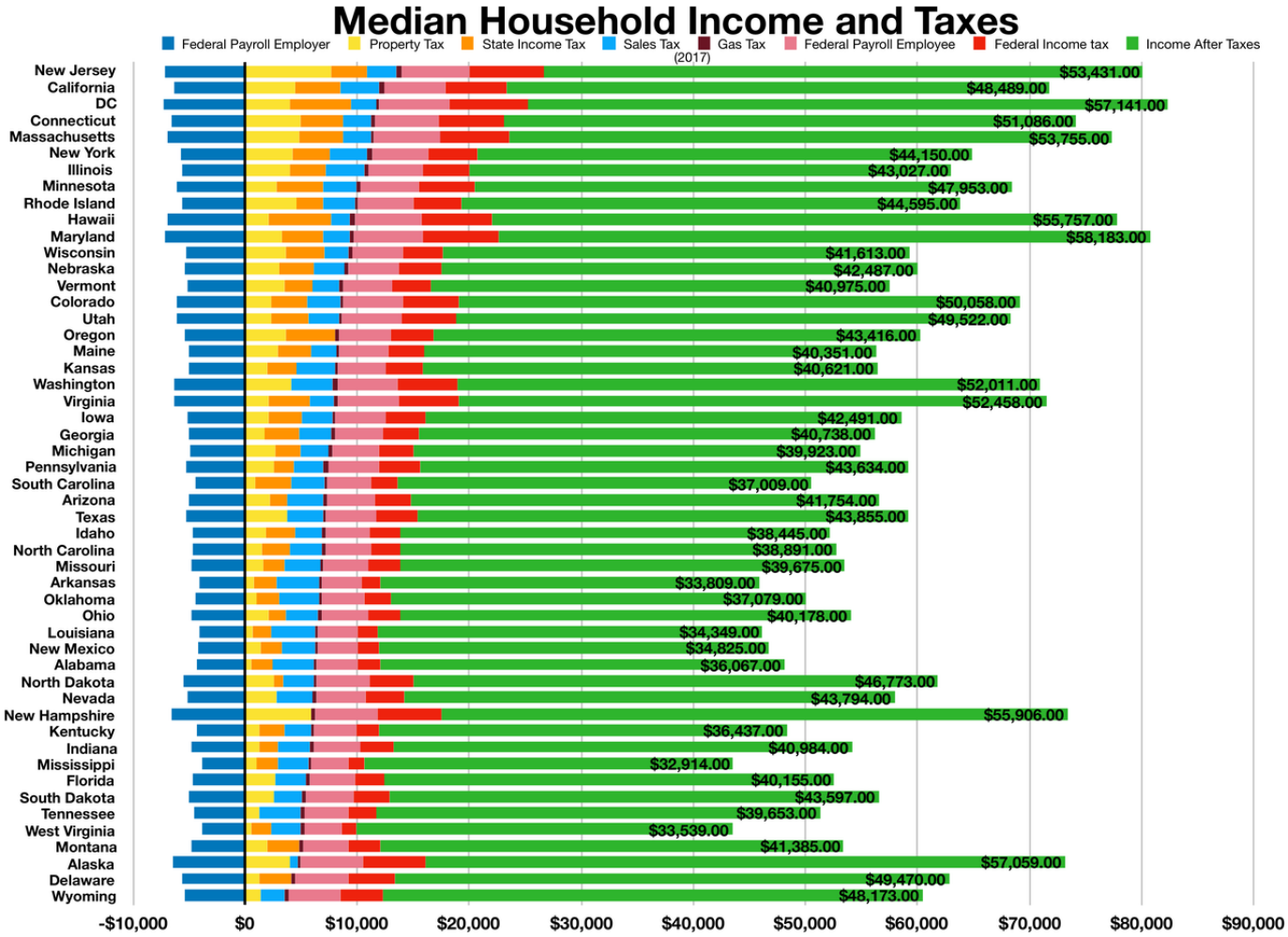
- Quantify energy, demand, emissions, and cost reductions for the Electric Power Board of Chattanooga, TN (EPB)



#	Description	Category	Value	Source
1	Insulate Roof	Envelope	R-16.12 to R-28.57	IECC-2012
2	Reduce Space Infiltration	Envelope	Reduce 25% from vintage	EnergyStar whole-house
3	Adjust Thermostat Setpoint (4F)	HVAC	4°F 2 hrs prior to peak	EPB
4	Smart Thermostat (8F)	HVAC	8°F 4 hrs prior to peak	EPB
5	Change Electric HVAC COP	HVAC	COP to 3.55 (ht) 3.2 (cl)	IECC-2012
6	Change Lighting Power Density	Lighting	LPD 0.85 W/ft^2	IECC-2012
7	Change to Gas Water Heater	Water	Efficiency 80% (assumes electric)	IECC-2012
8	Change to Gas HVAC	HVAC	Efficiency 80% (assumes electric)	IECC-2012

Tax data

- Tax Assessor's Data
  - Only 22 of 195 countries do not have a property tax
- U.S. Ad Velorem tax law
  - "according to value"



Source: https://en.wikipedia.org/wiki/Taxation_in_the_United_States

Tax data

- Overview of parcel data
  - 3,142 counties
  - No standardization of parcel data

CODE	DESCRIPTION
01	County Owned
02	City Owned
03	State Owned
04	Federally Owned
05	Religious Usage
06	Fraternal Usage
07	Office of State Assessed Properties
08	Commercial
09	Deletes or Combines
10	Industrial
11	Chattanooga Housing Authority
12	Back Tax
13	Hospital Authority
14	County Schools
15	City Schools
16	Cemeteries
17	Farms
18	Forest
19	Homestead
20	Mineral
21	Community Lot
22	Residential
23	Town of Lookout Mtn
24	Signal Mtn
25	East Ridge
26	Red Bank
27	Soddy Daisy
28	Collegedale
29	Ridgeside
30	Lakesite
31	Walden
32	Rental Property 40%, e.g. Apartment
33	Former Greenbelt
34	Open Space
35	City & County Owned
36	Airport Authority
37	EPB
40	Apartment (117)
45	Golf Course
50	Mobile Home Pk
51	Vacant MH Pad
97	Leasehold Assessment
98	In-Lieu of & Deferred Taxes
99	Unworked Parcel

Tax data

- Numeric code entries

Field	Starts	End	Length
OwnerName [1..3]	0017	0466	0450
PropAddress	0537	0569	0060
Parcel	0621	0632	0012
TaxMapNumber	0633	0647	0015
LotDim1	1082	1111	0030
LotDim2	1112	1126	0015
CalcAcres	1127	1146	0020
MapAcres	1147	1166	0020
DeedAcres	1167	1186	0020
LUCode	1187	1205	0019
NeighborhoodCode	1206	1217	0012
LandValue	1218	1237	0020
BuildingValue	1238	1257	0020
AppraisedValue	1258	1277	0020
AssessedValue	1278	1297	0020
District	1298	1301	0004
Zoning	1307	1310	0004
PropType	1330	1331	0002
ExemptCode	1332	1335	0004
SaleYear[1..4]	1566	1569	0004
SaleMonth[1..4]	1570	1571	0002
SaleDay[1..4]	1572	1573	0002
SaleConsideration[1..4]	1574	1592	0019
SaleType[1..4]	1633	1639	0007
SaleConf[1..4]	1640	1643	0004
Subdivision	1916	1935	0020
CurrentUse	1948	1951	0004

Code	District
1	City
2	County
3	County
2C	Collegedale
2E	East Ridge
2R	Ridgeside
3L	Lookout Mountain
3LS	Lakesite
3R	Red Bank
3S	Signal Mountain
3SD	Soddy Daisy
3W	Walden

Code	Definition
A	Assumption Deed
B	Combination Sale
C	Deed In Lieu of Foreclosure
D	Deed of Correction
E	Deed -Tenants by Entirety
F	Decree of Redemption
G	Divorce
H	Family Sale
I	Government Sale
J	Judgement
K	Partnership
L	Refile Deed
M	Sale Including Personal Property
N	Substitute Trustee Deed
O	Tax Sales
P	Trustees Sale
Q	Will
S	Quitclaim Deed
T	Tax Order
V	Charter
W	Warranty Deed
X	Master Deed
Y	Deed Reference

Tax data – building info!

CODE	DESCRIPTION
01	County Owned
02	City Owned
03	State Owned
04	Federally Owned
05	Religious Usage
06	Fraternal Usage
07	Office of State Assessed Properties
08	Commercial
09	Deletes or Combines
10	Industrial
11	Chattanooga Housing Authority
12	Back Tax
13	Hospital Authority
14	County Schools
15	City Schools
16	Cemeteries
17	Farms
18	Forest
19	Homestead
20	Mineral
21	Community Lot
22	Residential
23	Town of Lookout Mtn
24	Signal Mtn
25	East Ridge
26	Red Bank
27	Soddy Daisy
28	Collegedale
29	Ridgeside
30	Lakesite
31	Walden
32	Rental Property 40%, e.g. Apartment
33	Former Greenbelt
34	Open Space
35	City & County Owned
36	Airport Authority
37	EPB
40	Apartment (117)
45	Golf Course
50	Mobile Home Pk
51	Vacant MH Pad
97	Leasehold Assessment
98	In-Lieu of & Deferred Taxes
99	Unworked Parcel

Column	Field Description	Data Type
1	Parcel ID	Alpha
2	Exterior Type Code	Alpha
3	Exterior Type Description	Alpha
4	Jurist Code	Alpha
5	Jurist Description	Alpha
6	Year Built	Numeric
7	Taxable Building Amount	Numeric
8	Size Adjusted Area	Numeric
9	Story Height	Alpha
10	Roof Structure Code	Alpha
11	Roof Structure Description	Alpha
12	Roof Cover Code	Alpha
13	Roof Cover Description	Alpha
14	Prime Wall Code	Alpha
15	Prime Wall Description	Alpha
16	Second Wall Code	Alpha
17	Second Wall Description	Alpha
18	Heat Type Code	Alpha
19	Heat Type Description	Alpha
20	Account Number	Numeric
21	Card Number	Numeric
22	Street Number	Alpha
23	Street Name	Alpha
24	Land Use Code	Alpha
25	Land Use Description	Alpha
26	City	Alpha

Outline/Agenda

- Land Use Codes – residential

subdivision	division	category
100 RESIDENTIAL	110 Household Units	111 One Family Household Unit
		112 Two Family Household Unit (Duplex)
		113 Multi-Family (4-9 unit Apartment)
		114 Two Family Units (Duplex Owner Occupied)
		115 Triplex
		116 Condominium,
		117 Apartment: 10 units or more
	120 Group Quarters	Rooming and boarding houses, fraternity and sorority house and other membership lodgings. Nursing homes, college dormitories, other halls or dormitories, retirement homes, religious quarters, orphanages, convents.
	130 (Reserved for future use)	
	140 Mobile Homes	141 Mobile Homes (Single Trailer)
		142 Mobile Home Park
		143 Mobile Home Park (Privately Owned)
	150 Transient Lodging	Motels, tourist courts, lodges, hotels and other transient lodgings.

Outline/Agenda

- Land Use Codes – commercial

500 WHOLESALE AND RETAIL TRADE	530 Retail Trade - General	531 Department stores
		536 Discount department stores
		537 Antiques and second hand
		538 Drugstores
	540 Retail Trade - Food	541 Groceries (supermarkets)
		542 Meats and fish markets
		543 Fruits and vegetables
		544 Candy, nuts and confectionery
		545 Dairy products
		546 Bakeries
		547 Liquor
		548 Groceries, convenience shops (drive-in type)

400 TRANSPORTATION, COMMUNICATIONS AND UTILITIES	490 Warehouse	491 (Reserved for future use)
		492 Warehouse Storage
		493 Distribution-Warehouse

500 WHOLESALE AND RETAIL TRADE	510 Wholesale Trade	Automotive equipment, drugs, chemicals, dry goods and apparel, groceries, farm products, electrical goods, hardware, plumbing, heating equipment and supplies, machinery, equipment and supplies.
	520 Retail Trade Equipment	521 Lumber and other building materials
		522 Heating and plumbing equipment
		523 Paint, glass and wall paper
		524 Electrical supplies
		525 Hardware and farm equipment

Outline/Agenda

- Land Use Codes – commercial

	570 Retail Furniture, Home Furnishings and Equipment	Furniture and home furnishings, household appliances, radios, televisions and home electronic supplies and music supplies, florists and garden supplies, floor coverings, draperies, china, glass and metal ware.
	580 Retail Trade Eating and Drinking	583 Restaurants typically those which provide full-course meals
		584 Diners and luncheonettes characterized by counter service, limited.
		585 Snack bars, drive-ins with window and/or car service, possibly limited counter service.
		586 (Reserved for future use)
		587 Bars and taverns
		590 Other Retail Trades
600 SERVICES	680 Educational Services	682 University, colleges and junior colleges
		683 Vocational and special training
		684 Nursery schools and day care centers
		685 Elementary schools
		686 Junior high schools
		687 Senior high schools
	690 Other Services	691 Churches-, synagogues, and temples

Parcel ID	ET Code	Exterior Type Description	J Code	Jurist Description	Year Built	Taxable Bldg Amnt	Size Adjusted Area	Story Height	RS Code	Roof Structure Description	R Code	Roof Cover Description	P C
Redact	2	DUPLEX			1975	0	2317	1	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R
Redact	105	2 STORY			1973	0	2156	2	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R
Redact	103	SPLT LVL/FOY	111	ONE FAMILY U	1970	0	1400	1	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R
Redact	102	RANCH			1969	0	1951	1	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R
Redact	2	DUPLEX	112	TWO FAMILY U	1971	0	1812	1	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R

Outline/Agenda

- Building-specific data

Parcel ID	ET Code	Exterior Type Description	J Code	Jurist Description	Year Built	Taxable Bldg Amnt	Size Adjusted Area	Story Height	RS Code	Roof Structure Description	R Code	Roof Cover Description	PW Code
Redact	2	DUPLEX			1975	0	2317	1	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R_11
Redact	105	2 STORY			1973	0	2156	2	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R_12
Redact	103	SPLT LVL/FOY	111	ONE FAMILY U	1970	0	1400	1	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R_33
Redact	102	RANCH			1969	0	1951	1	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R_12
Redact	2	DUPLEX	112	TWO FAMILY U	1971	0	1812	1	R_02	HIP/GABLE	R_04	SHINGLE ASPH	R_12

Prime Wall Description	SW Code	Second Wall Description	HT Code	Heat Type Description	Account#	Card#	Street#	Street Name	LU code	Land Use Description	City
WOOD FR W SH	R_12	BRICK	R_21	CENTRL HEAT&	Redact	1	Redact	Redact	DU	DUPLEX-OWN O	HIXSON
BRICK	R_11	WOOD FR W SH	R_21	CENTRL HEAT&	Redact	1	Redact	Redact	RS	RESIDENTIAL	HIXSON
VINYL			R_21	CENTRL HEAT&	Redact	1	Redact	Redact	RS	RESIDENTIAL	HIXSON
BRICK			R_21	CENTRL HEAT&	Redact	1	Redact	Redact	RS	RESIDENTIAL	HIXSON
BRICK			R_21	CENTRL HEAT&	Redact	1	Redact	Redact	MF	MULTI-FAMILY	HIXSON

Outline/Agenda

- Overview of files (for this example)
  - GIS parcel data (*.shp)
  - Parcel data (*.zip)
  - Letter (*.pdf) – instructions for DOS prompt executable to extract 300MB of data into CSV
- Summary of tax assessor data
  - Useful source of building-specific data
    - Building type, Height, Size, Envelope type (Wall), HVAC, Roof type, Fenestration (sometimes), Footprints (often just parcel, not building)
  - Limitations
    - County-specific data encoding, data format (PDF>Excel), public vs. non-public information, undocumented processes for acquiring/using/sharing data

Results – Tax Assessor

- Statistical summary of 1 county

Size Adjusted Area	%	Story Height	%	Land Use code	%	Heat Type Description	%
1,000-1,499	32.6	1	74.5	RESID	82.3	CENTRL HEAT&	75.4
1,500-1,999	22.1	2	17.4	COMM	7.9	<EMPTY>	12.6
2,000-2,499	12.2	1.5	7.1	MFG	4.7	GRAVITY	7.4
5,000+	9.8	3	0.6	IN	2.4	NO HVAC	3.4
500-999	8.3	>7	0.6	AG	1	FORCED HOT A	0.9
2,500-2,999	6.6	2.5	0.1	EX	0.8	GHA	0.1
3,000-3,499	3.8	4	0.1	DU	0.4	CENTRAL A/C	0.1
3,500-3,999	2.1	5	0.0	EID	0.2	REV CYCLE UN	0.0
4,000-4,499	1.3	6	0.0	RLS	0.1	CENT HEAT &	0.0
4,500-4,999	0.8	7	0.0	BCMT	0.1	NONE	0.0
Roof Structure Description	**%**	**Roof Cover Description**	**%**	**Decade**	**%**	**Prime/Second Wall Description**	**%**
HIP/GABLE	86.0	SHINGLE ASPH	81.9	2000	13.7	<EMPTY>	41.9
WOOD RAFTERS	2.8	SHEET METAL	3.5	1960	13.6	WOOD FR W SH	15.5
BAR JOISTS	2.4	BUILT-UP	3.4	1970	13.5	VINYL	13.5
OPEN STEEL S	2.1	METAL	2.9	1990	11.6	BRICK	13.1
STEEL TRUSS	1.8	<EMPTY>	1.7	1950	11.4	WOOD FR ASBT	2.7
<EMPTY>	1.7	ASPHALT SHIN	1.7	1980	10.8	CONC BLK PLA	1.6
NONE	0.8	CORRUGATED M	1.3	1940	7.6	ALUMINUM	1.4
WOOD TRUSS	0.6	NONE	0.9	2010	6.9	HARDIE BOARD	1.4
FLAT/SHED	0.5	ROLL COMP	0.6	1930	4.8	BRICK VENEER	1.2
GAMBREL	0.3	BUILT UP T &	0.4	1920	3.8	CORRUGATED M	1.2

Outline/Agenda

- Context
- Tax Assessor's Data
- Comparison Matrix
- Scalable Data Sources

Sensitivity analysis

- Sensitivity analysis for all building types
  - 80% of commercial buildings - 16 climate zones, 16 building types, averaging 5.75 vintages
  - 281-4,617 building descriptors (e.g. thermostat, insulation level) were modified

	Small Office	Outpatient	Large Office	Medium Office	Hospital	Warehouse	Small Hotel	Large hotel
Inputs	458	3483	1072	760	1955	333	1823	887
	Strip Mall	Retail	Quick Service Restaurant	Full Service Restaurant	Mid Rise Apt	High Rise Apt	Secondary School	Primary School
Inputs	800	438	281	286	1464	4617	1621	1051

  - Fractional Factorial (FrF2) resolution IV statistical design of experiments
- Summarize 768 lists of impactful variables
  - 254,544 annual simulations were completed on the nation's fastest supercomputer (Titan)
  - 216 Excel spreadsheets were created listing the energy and demand impacts of each building property
- Quantify Most Important Building Parameters
  - Top 10 annual energy (kWh) and demand/peak-shaving (kW) variables for each of the 16 building types

Comparison Matrix - example

- 6 types of data, 37 sources of data, how do you compare?
  - Satellite and airborne imagery
  - Cartographic data
  - Ground level images
  - Elevation data
  - Building information databases
  - 3D building model databases

	Short Title
Summary	Satellite imagery, including panchromatic and multispectral images
Data type	Image
Company	
Website	
Temporal resolution	Cities - 3-11 times per week
Spatial resolution	0.3 m
Measure accuracy	
Cost	$11 per sq. km
Format	GeoTiff
Mapping to building input variables	Building footprints
Mapping to area properties	Vegetated areas, road surface, buildings, parking lots
Mapping to material properties	Road pavement materials (e.g., concrete, asphalt), parking lots (e.g., gravel, soil)
Coverage of US	Over 10 million km^2 of coverage of the contiguous US
Orientation	Aerial
Existing internal software	N/A
Existing expertise	Remote sensing data analysis tool
Restrictions	N/A
Comments	

Comparison Matrix – fields (1 of 2)

- Potentially useful fields to consider when comparing data sources

1. Title -- short label for referring to the dataset
2. Summary -- short description of the data
3. Data type -- the format in which the information is stored (usually image, database, or computationally derived from multiple data sources)
4. Company -- name of the organization that makes the data available
5. Website -- hyperlink to the most pertinent information necessary for using this dataset
6. Temporal resolution -- how often the datasets are collected (e.g., 2–5 years)
7. Spatial resolution -- the dimensions of the data (e.g., 1 km^2 per pixel)
8. Measure accuracy -- information available regarding the accuracy of the database based on input sources or sensor calibration
9. Cost -- any initial or recurring costs required to access/retain rights to the data
10. Format -- the standard file format in which the datasets are stored
11. Mapping to building input variables -- for models (e.g. building type, square footage, window-to-wall ratio, façade material type, thickness, density)

Comparison Matrix – fields (2 of 2)

- Potentially useful fields to consider when comparing data sources

12. Mapping to area properties -- indicates whether these datasets are useful in segmenting area type (e.g., buildings, roads, open/vegetated spaces)
13. Mapping to material properties -- indicates whether these datasets are useful in determining material types (e.g., concrete, brick, soil, gravel, asphalt, granite)
14. Coverage of United States (US) -- indicates the extent to which the data provided are local versus national
15. Orientation -- where relevant, the general view from which the data were taken (e.g., street view, single side of a building, multiple sides of building, perspective, oblique)
16. Existing internal software -- does the current team have software capabilities that leverage this dataset for purposes that could be synergistically leveraged for this project
17. Existing expertise -- does the current team have any unique knowledge or skills that would be vital to the successful use of the data for this project
18. Restrictions -- what are the limitations on the use of the data (e.g., legal/privacy ratings, number of Application Program Interface [API] calls per day)
19. Comments -- any major observations about the data that do not fit in the previous categories

Outline/Agenda

- Context
- Tax Assessor's Data
- Comparison Matrix
- Scalable Data Sources

Scalable Data – example 1

	DigitalGlobe Standard Imagery	DigitalGlobe Precision Aerial image
Summary	Satellite imagery including panchromatic and multispectral images (4 bands or 8 bands)	Aerial imagery, including panchromatic and multispectral images
Data type	Image	Image
Company	DigitalGlobe	DigitalGlobe
Website	www.digitalglobe.com	www.digitalglobe.com
Temporal resolution	N/A	N/A
Spatial resolution	Pan: 0.5/0.6 m; MS 2.0/2.4 m	0.3 m
Measure accuracy	High	High
Cost	Pan: $24 per sq. km; Pan+MS $27 per sq. km	$11 per sq. km; Pricing URL
Format	GeoTiff	GeoTiff
Building inputs	Building footprint	Building footprints
Area properties	Vegetated areas, road surface, buildings, parking lots	Vegetated areas, road surface, buildings, parking lots
Material properties	Road pavement materials (e.g., concrete, asphalt), parking lots (e.g., gravel, soil)	Road pavement materials (e.g., concrete, asphalt), parking lots (e.g., gravel, soil)
Coverage of US	High	Over 10 million km^2 of coverage of the contiguous US
Orientation	Aerial	Aerial
Existing expertise	Remote sensing data analysis tool	Remote sensing data analysis tool
Restrictions	Contract-specific	Contract-specific

Aerial view - footprints

Algorithm: Deep Learning extended and using GPUs for fast building footprint and area extraction over large geographical areas.

Multi-company Competition Precision/Recall – 30/35; Current Precision/Recall – 60+/60+

Scalable Data – example 2

	National Elevation Dataset
Summary	Ground elevation data
Data type	Raster
Company	USGS
Website	http://ned.usgs.gov/
Temporal resolution	N/A
Spatial resolution	1/3, 1, and 2 seconds of arc; 1/9 arc-second and 1 meter for some areas
Measure accuracy	Mean square error is 1.55 m
Cost	Free
Format	Raster data
Building inputs	Main floor ground elevation
Area properties	Road surface elevation
Material properties	N/A
Coverage of US	High
Orientation	Aerial
Existing expertise	GIS software
Restrictions	Restrictions URL

Summary

Data type

Company

Website

Temporal resolution

Spatial resolution

Measure accuracy

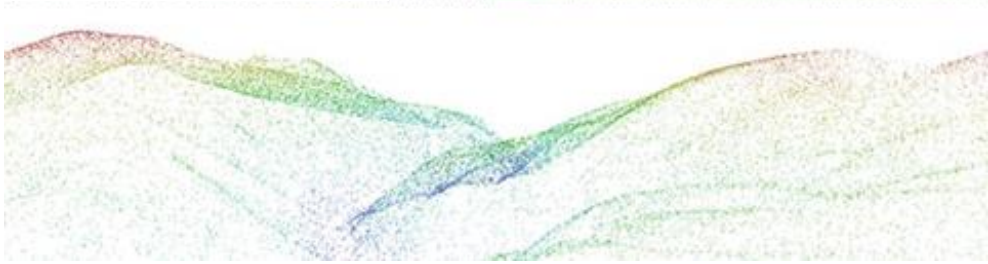
Cost

Format

Building inputs

Area properties

Material properties

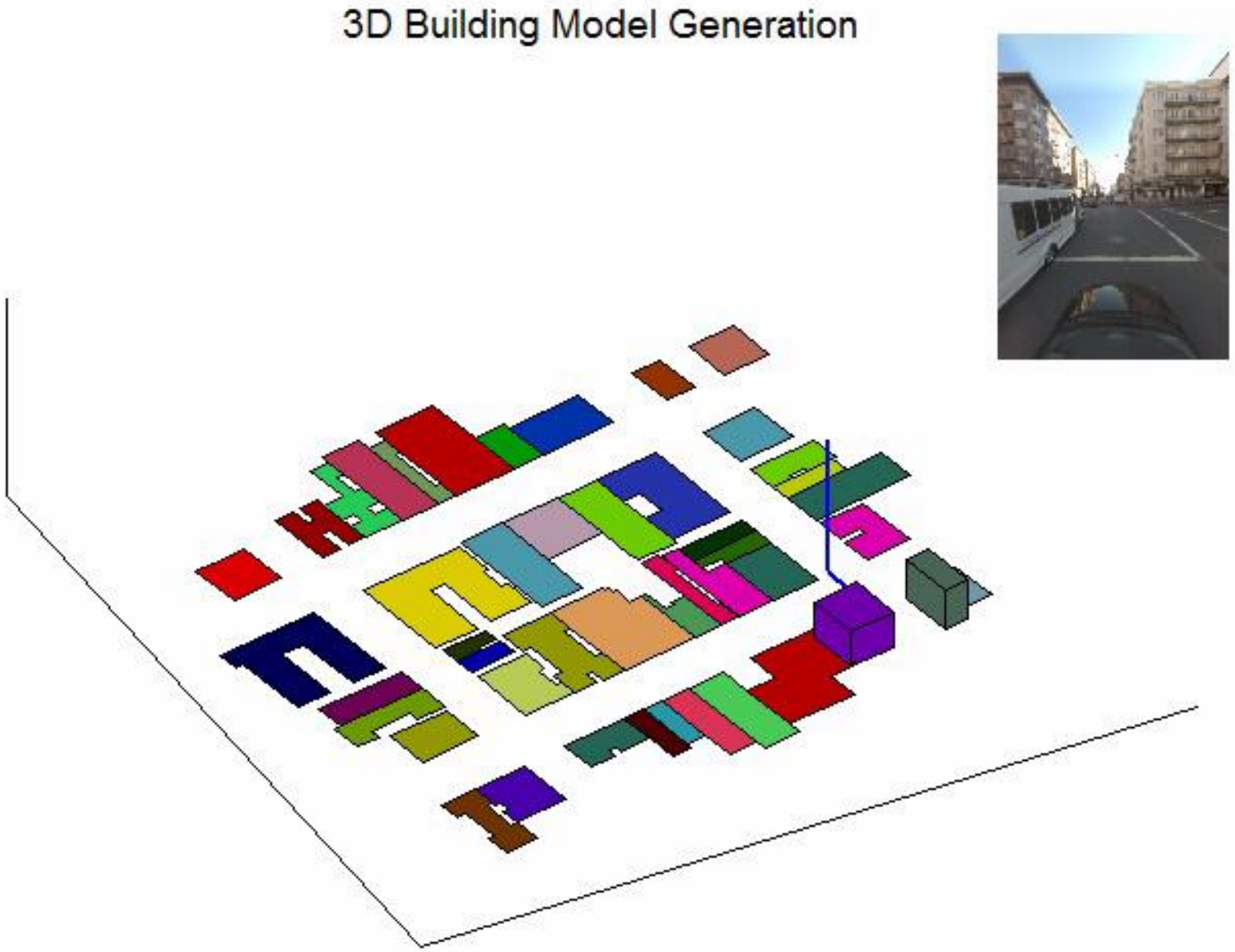


LiDAR – building heights

Scalable Data – example 3

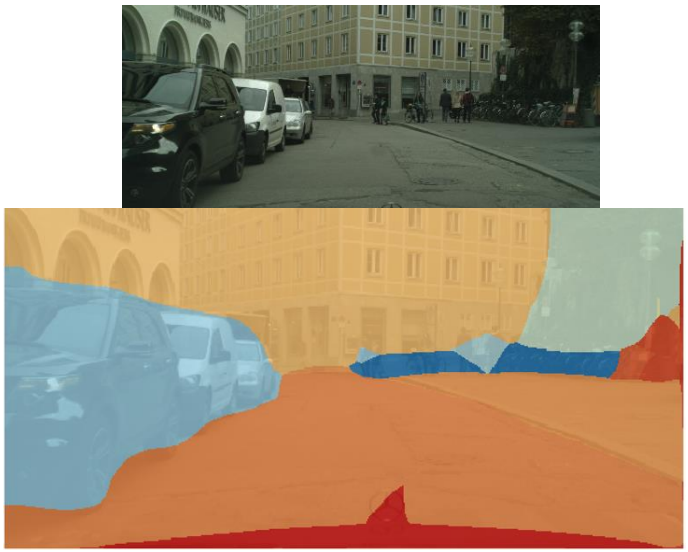
	Google Street View
Summary	Street view images. Downloadable using Google Street View API
Data type	Image
Company	Google
Website	Developer URL
Temporal resolution	N/A
Spatial resolution	N/A
Measure accuracy	Location errors exist
Cost	Free
Format	jpg
Building inputs	Height, window-to-wall ratio
Area properties	N/A
Material properties	Road pavement materials (e.g., concrete, asphalt), building exterior materials (e.g., glass, concrete), parking lots (e.g., gravel, soil)
Coverage of US	High
Orientation	Multi-side
Existing internal software	Building height estimation
Existing expertise	OpenCV
Restrictions	25,000 API calls per day. Restrictions URL

StreetView - height

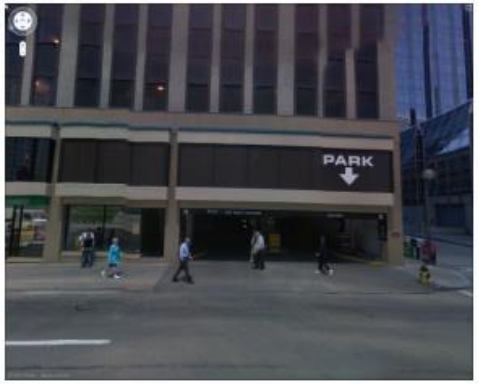


StreetView – Façade, WWR

Façade Type

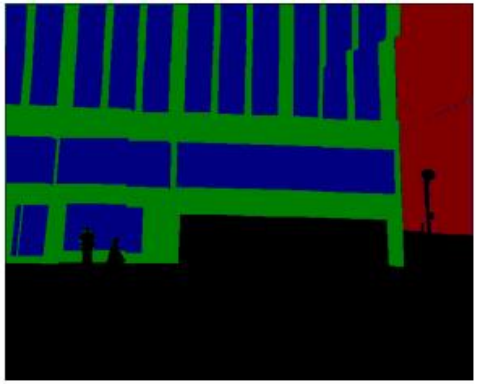


Windows (blue)
Façade (green)
Street/open (black)
Other building (red)

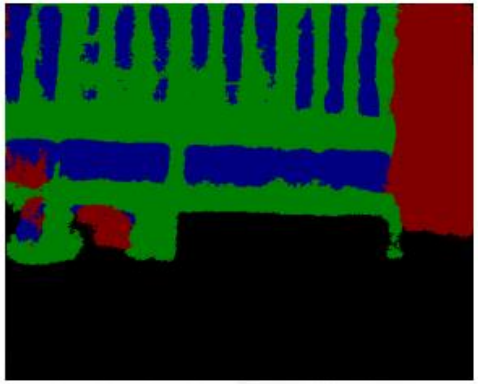


Input image

Window-to-wall ratio



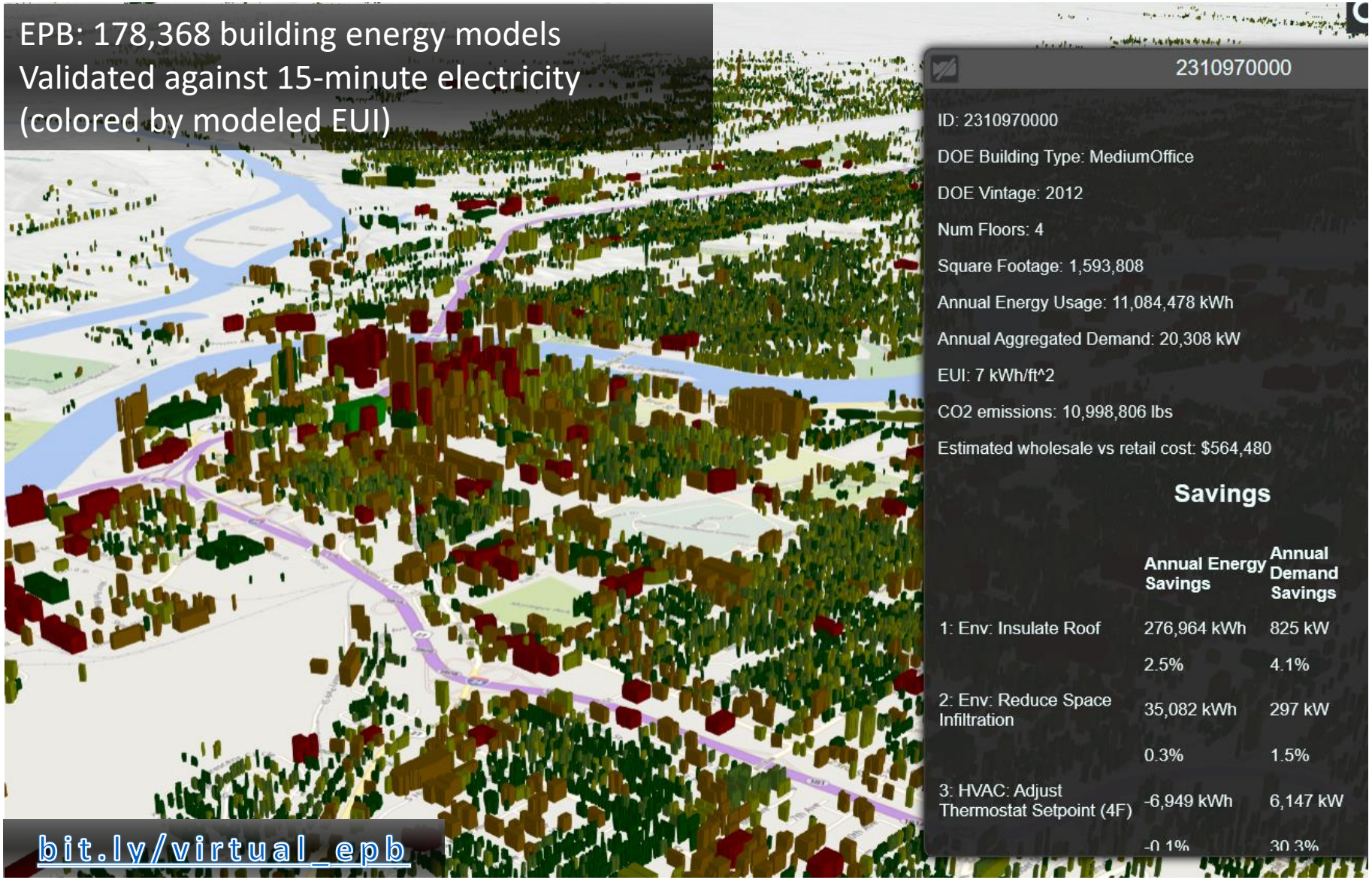
Ground truth



Model output

Results – Digital Twin of a Utility

EPB: 178,368 building energy models
Validated against 15-minute electricity
(colored by modeled EUI)



bit.ly/virtual_epb

Bibliography

Yuan, Jiangye, Joshua R. New, Jibonananda Sanyal, and Olufemi Omitaomu. 2015. “Urban Search Data Sources.” ORNL internal report ORNL/TM2015/397.

Copeland, William, and Joshua R. New. 2019. “Digital Twin of a City Utility: Issues, science, implementation, and results.” Proceedings of the Better Buildings Summit, Data of the Future: Digital Cities.

New, Joshua New. 2018. Automatic Building Detection and Energy Model Creation (AutoBEM). Presented at the 2018 Building Performance Analysis Conference, Atlanta, Georgia, September 26-28.

New, Joshua R., Adams, Mark, Im, Piljae, Yang, Hsiuhan, Hambrick, Joshua, Copeland, William, Bruce, Lilian, Ingraham, James A. (2018). "Automatic Building Energy Model Creation (AutoBEM) for Urban-Scale Energy Modeling and Assessment of Value Propositions for Electric Utilities." In *Proceedings of the International Conference on Energy Engineering and Smart Grids (ESG)*, Fitzwilliam College, University of Cambridge, Cambridge city, United Kingdom, June 25-26, 2018.

New, Joshua R., Bhandari, Mahabir, Shrestha, Som, and Allen, Melissa. (2018). "Creating a Virtual Utility District: Assessing Quality and Building Energy Impacts of Microclimate Simulations." In *Proceedings of the International Conference on Sustainable Energy and Environmental Sensing (SEES)*, Cambridge, UK, June 18-19, 2018.

Ingraham, James A. and New, Joshua R. (2018). "Virtual EPB." Presented to Building Technologies Office following the *BTO Peer Review*, 87 slides. Arlington, VA, May 3, 2018.

AutoGen – Automatic EnergyPlus file modifier/Generator, worlds fastest building energy model creator utilizing text replacement for variable in EnergyPlus files; awarded by U.S. Copyright Office under registration number TXu 2-159-000.

AutoSim – Automatic Simulator, (CR17-00072, UTB80000011) - worlds fastest buildings simulator for scalably distributing EnergyPlus files on High Performance Computing devices, simulating on virtual disk, and returning results for storage and analysis; awarded by U.S. Copyright Office under registration number TXu 2-141-960.

QUESTIONS?

Joshua New

newjr@ornl.gov